2020

Voxel carving-based 3D reconstruction of sorghum identifies genetic determinants of light interception efficiency

Mathieu Gaillard

Chenyong Miao

James Schnable

Bedrich Benes

Received: 7 July 2020 | Accepted: 9 July 2020

DOI: 10.1002/pld3.255

ORIGINAL RESEARCH

Voxel carving-based 3D reconstruction of sorghum identifies genetic determinants of light interception efficiency

Mathieu Gaillard[1] | Chenyong Miao[2] | James C. Schnable[2] | Bedrich Benes[1,3]

[1]Department of Computer Graphics Technology, Purdue University, West Lafayette, IN, USA

[2]Center for Plant Science Innovation and Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, USA

[3]Department of Computer Science, Purdue University, West Lafayette, IN, USA

Correspondence
James C. Schnable, Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68588, USA.
Bedrich Benes, Computer Science, Purdue University, West Lafayette, IN 47906, USA.
Email: schnable@unl.edu (JCS), bbenes@purdue.edu (BB)

Funding information
Research reported in the publication was supported by the Foundation for Food and Agriculture Research (award number—Grant ID: 602757) to BB and JCS. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the FFAR. This research was supported by the Office of Science (BER), U.S. Department of Energy, Grant no. DE-SC0020355 to JCS

Abstract

Changes in canopy architecture traits have been shown to contribute to yield increases. Optimizing both light interception and light interception efficiency of agricultural crop canopies will be essential to meeting the growing food needs. Canopy architecture is inherently three-dimensional (3D), but many approaches to measuring canopy architecture component traits treat the canopy as a two-dimensional (2D) structure to make large scale measurement, selective breeding, and gene identification logistically feasible. We develop a high throughput voxel carving strategy to reconstruct 3D representations of sorghum from a small number of RGB photos. Our approach builds on the voxel carving algorithm to allow for fully automatic reconstruction of hundreds of plants. It was employed to generate 3D reconstructions of individual plants within a sorghum association population at the late vegetative stage of development. Light interception parameters estimated from these reconstructions enabled the identification of known and previously unreported loci controlling light interception efficiency in sorghum. The approach is generalizable and scalable, and it enables 3D reconstructions from existing plant high throughput phenotyping datasets. We also propose a set of best practices to increase 3D reconstructions' accuracy.

KEYWORDS
3D plant reconstruction, phenotyping, quantitative genetics, sorghum

1 | INTRODUCTION

Yields of agricultural plants will need to grow 70% by 2050 to meet increasing demands for food and feed, which are projected to double between 2005 and 2050 (Alexandratos & Bruinsma, 2012; Tilman, Balzer, Hill, & Befort, 2011). Currently, we are not on track to meet this demand. In many parts of the world, increases in wheat and rice yields have dropped to zero, as observed by Grassini, Eskridge, and Cassman (2013). While maize yields continue to increase, the increased annual spending on breeding is required each year to achieve the same fixed annual increase in yields per year (Grassini et al., 2013). Developing new high yielding and more resilient crop varieties depends on the ability to score large populations of new lines for traits. All else constant, the more unique data can be

Mathieu Gaillard and Chenyong Miao contributed equally to this work.

This manuscript has been previously deposited as a preprint at https://doi.org/10.1101/2020.04.06.028605.

collected from, and the more accurate that data, the faster the rate of genetic gain within a breeding program. Unfortunately, the cost of collecting trait data from crop varieties has remained constant or increased per data point. One way to improve this process's efficiency is to focus on better, more precise, and faster extraction of meaningful information from the collected data.

High throughput phenotyping technology is an umbrella term that describes a wide range of new approaches that leverage advances in engineering and computer science that seek to address this current bottleneck. Generally, new high throughput phenotyping technologies are approaches to measuring plant phenotypes that are predicted to be either (a) lower cost per data point, (b) more accurate, or (c) enable the measurement of yield relevant traits which it is not presently feasible to score.

Early infrastructure investments in high throughput phenotyping focused on automated data acquisition in controlled environment plant growth facilities (Fahlgren et al., 2015; Ge, Bai, Stoerger, & Schnable, 2016; Junker et al., 2015). In these facilities plants are moved around in conveyor belts (Figure 1a) and on regular intervals are brought into a series of imaging chambers (Figure 1b) where they are photographed from several angles using different types of cameras (see Figure 1d). Several software tools have been developed to extract different phenotypic data from the individual two-dimensional (2D) images generated by these automated controlled environment phenotyping facilities (Gehan et al., 2017; Lobet, 2017).

One essential set of features linked to yield, which are difficult to quantitate accurately from 2D images, are traits related to canopy architecture, including leaf number, leaf angle, leaf length, internode spacing, etc. Within-species variation in light interception efficiency, water use efficiency, and crop yield have all been linked to differences in canopy architecture (Hammer et al., 2009; Maddonni, Chelle, Drouet, & Andrieu, 2001; Westgate, Forcella, Reicosky, & Somsen, 1997). A large proportion of the yield gain per unit area over the past half-century has come from increased planting density and breeding for lines that can thrive at higher planting densities (Duvick, 2005). In maize, selection for yield at high density indirectly selected for lines with more vertical leave angles, spreading the same amount of incident light over a sizeable photosynthetic surface (Duvick, 2005; Pendleton, Smith, Winter, & Johnston, 1968; Pepper, Pearce, & Mock, 1977). In sorghum, breeders have selected for large-effect mutations with reduce internode spacing, producing a denser canopy closer to the ground (Quinby et al., 1953). Future efforts to engineer efficient canopy architectures will require quantitation and simulation of a range of canopy architectures achievable from natural genetic variation in crop species (Benes et al., 2020; Marshall-Colon et al., 2017) Collecting the data needed for this objective, in turn, requires a more comprehensive collection and characterization of genotype to genotype variation in three-dimensional (3D) canopy architecture on a large scale.

However, plants are complex 3D structures, and data collected from a single or several 2D images can miss or inaccurately estimate for important plant features (McCormick, Truong, & Mullet, 2016; Thapa, Zhu, Walia, Yu, & Ge, 2018). Various approaches attempt to faithfully extract either the full 3D structure of plants or at least some essential traits from captured data. LIDAR scanning using a sensor mounted on a robotic arm has been used to reconstruct 3D models of barley plants, achieving $R^2 = .96$ with ground truth measurements to predict the area of individual leaves (Paulus, Schumann, Kuhlmann, & Léon, 2014). LIDAR-based reconstruction of maize and sorghum plants using a fixed LIDAR sensor and plants were placed on a rotating platform achieved $0.92 \leq R^2 \leq .94$ for maize and sorghum (Thapa et al., 2018). Time-of-flight cameras (e.g., Microsoft Kinect) were employed to generate 3D models of individual plants from a sorghum RIL population, achieving $R^2 = .85$ with destructively measured leaf area (McCormick et al., 2016). However, these approaches require dedicated equipment unlikely to be available in many plant research labs. LIDAR is not suitable for data with high-frequency information and high inter reflections that are typical for vegetation. Small parts can be easily missed, and the reconstruction algorithms often fail on small or slim parts such as leaves. Conventional RGB cameras are easily accessible, and some faculties around the world use conveyor belts and rotating platforms to image individual plants from multiple angles at regular intervals (Fahlgren et al., 2015; Ge et al., 2016; Junker et al., 2015; Yang et al., 2014).

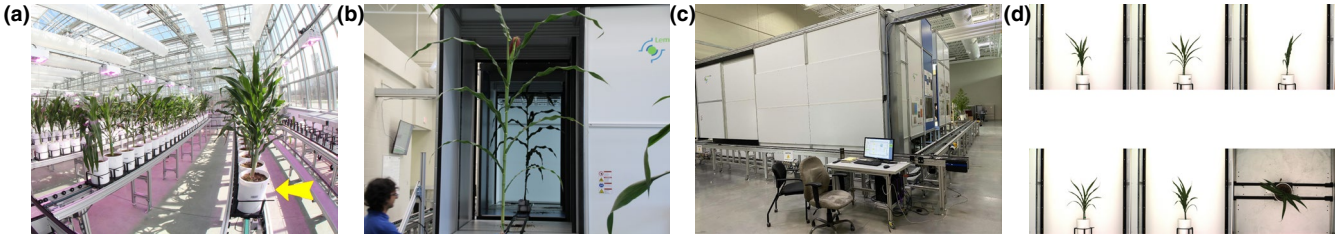

FIGURE 1 University of Nebraska Greenhouse Innovation Center phenotyping facility: (a) The greenhouse fitted with conveyor belts where plants are stored and moved in an automated fashion to be watered and taken through imaging chambers. A plant on the conveyor belt is marked with an arrow. Plants shown in this image are members of the same Sorghum Association Panel analyzed in this paper. (b) A maize plant exiting the series of four imaging chambers employed at this faculty to capture different types of data. New maize plants in earlier imaging chambers are also visible in this image. (c) The control unit, a computer managing movement, imaging, water, and climate controls for the entire facility. (d) An example of the image data employed in this study. Here a single sorghum plant has been imaged from five different side view angles and a single photo was taken by the system from directly above the plant

One of the most commonly used method is Structure From Motion (Lou, Liu, Sheng, Han, & Doonan, 2014; Tomasi & Kanade, 1992; Quan et al., 2006) (SFM) that requires a large number of images (hundreds), long processing times (minutes to hours) and often fails on texture-less or highly specular surfaces and on data that include high-frequency noise, such as foliage. Although significantly more accessible than LIDAR, current approaches to 3D reconstruction from RGB images share the reconstruction problems from LIDAR data. Other approaches to 3D reconstruction attempt to generate individual parts of plants: for example, generalized cylinders (Tan, Zeng, Wang, Kang, & Quan, 2007; Tan, Fang, Xiao, Zhao, & Quan, 2008), skeletons (Du, Lindenbergh, Ledoux, Stoter, & Nan, 2019), or use optimization to find a generative model with a database of leaves (Ward et al., 2015), and even reconstruct the entire growth model (St'ava et al., 2014). However, these methods usually require human interaction, comprehensive input data, clear background, or other requirements that make them difficult to apply in phenotyping where hundreds of plants are scanned over long periods.

Another well-established method is space carving of Kutulakos and Seitz (2000) that reconstructs 3D voxels occupied by the captured plant. Its main advantage is that it requires fewer images and a lower processing time than SFM. The drawback is that the algorithm needs an exact calibration and segmentation of the object to reconstruct, whereas SFM estimates calibration automatically by matching key points between views. Space carving is suitable for phenotyping facilities because the environment is controlled, which eases calibration and segmentation. Recent contributions focus on seedlings, which are smaller plants that are easier to reconstruct (Golbach, Kootstra, Damjanovic, Otten, & Zedde, 2016; Koenderink et al., 2009; Klodt & Cremers, 2015). Other contributions focus on accelerating voxel carving with octrees (Scharr et al., 2017). Roussel, Geiger, Fischbach, Jahnke, and Scharr (2016) focuses on seeds as small as 200 μm, and they set up an experiment to evaluate the accuracy concerning the number of images used for reconstruction. Cabrera-Bosquet et al. (2016) reconstruct thousands of Maize plants using voxel carving on 13 views. Reconstructed plants are then used to estimate traits such as: light interception and radiation use efficiency. Regarding 3D reconstruction, one drawback is that the algorithm does not output a surface but the photo hull of the plant, which is the maximal photo-consistent volume in which the actual plant is contained. Automatic measurements are not straightforward with a voxel representation (Golbach et al., 2016). For a broader view on plant reconstruction, we refer the reader to the survey by Gibbs et al. (2017) about plant reconstruction.

Calibration and segmentation are crucial to the success of space carving. For calibration, the work of Li, Heng, Koser, and Pollefeys (2013) is of particular interest, because it enables calibrating multiple cameras with non-overlapping views, which is the usual case in rotating imaging platforms. As for the segmentation of plant pixels in 2D images, it can take several approaches. One such method is "differencing" when the image is compared to a reference image taken by the same camera in the same imaging chamber without a plant present (Choudhury, Samal, & Awada, 2019). A second widely adopted approach is excess green thresholding (Gehan et al., 2017). Supervised classification algorithms that consider both pixel RGB values and data from immediately adjacent pixels have been shown to exhibit higher accuracy in plant pixel segmentation than many thresholding or "differencing"-based methods (Adams, Qiu, Xu, & Schnable, 2020). For example, the work of Donné et al. (2016) employs a convolutional neural network to perform a segmentation.

Our contribution is to validate a 3D reconstruction pipeline for a very low number of images: only five sides and one top photos. Our approach extends the space carving algorithm to allow for a fully automatic generation of 3D voxel grids approximating the shape of the large numbers of scanned plants (Figure 2h). This approach builds on voxel carving, but aims to achieve fully automatic reconstruction for large numbers of plants and does not require any user interaction on a per plant basis. This scalability and lack of required human interaction enabled us to evaluate hundreds of plants' method, more than an order of magnitude higher than many previous studies. We benchmarked 3D extraction and reconstruction as taking less than one minute per plant to create a 3D voxel grid with a resolution 512^3 on a workstation equipped with an Intel Xeon W-2145 (8 cores at 3.7 GHz), with significant decreases in per plant processing time for processing multiple plants in series. We extended the voxel carving algorithm to favor recall over precision, and thus get smoother plant shapes for a subsequent skeletonization, as explained in our following paper (Gaillard, Miao, Schnable, & Benes, 2020). We studied the effect of the camera setup on the accuracy of the reconstruction of procedurally generated 3D maize plant models. We showed that

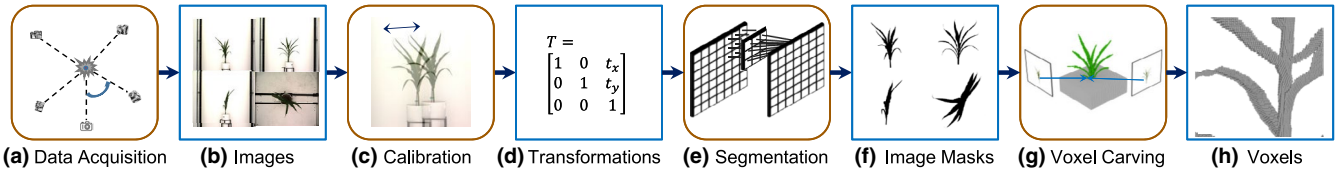

FIGURE 2 Overview of the 3D reconstruction process employed in this study. Round boxes indicate processes and squared boxes indicate data. (a) Images of the plant are acquired by the phenotyping chamber. (b) These images are precisely calibrated in an automated fashion to ensure the (c) the plant is properly centered in each of the images taken from different angles. The resulting calibrated images are (d) segmented (e) into binary masks distinguishing which pixels are plant (black) and not-plant (white). (f) Segmentation is performed by a convolutional neural net trained on a subset of manually segmented images. (g) These calibrated binary masks are employed as the input for the voxel carving process itself. (h) The output of voxel carving is the coordinates of a set of voxels corresponding to the three-dimensional space occupied by the target plant

some configurations are more efficient for the same investment in money. We evaluate the impact of an imprecise calibration on the resulting 3D plants. Finally, we showed that a 3D representation brings substantial information compared to only 2D images. Our algorithm allows for quantitation of traits light interception efficiency. Scoring large numbers of plants enabled quantitative genetic approaches to identify loci controlling variation in 3D traits. A number of inherently 3D traits were estimated from 3D reconstructions from 336 sorghum plants. These trait values, combined with information on hundreds of thousands of genetic markers, were used to determine the proportion of phenotypic variance for each trait controlled by genetic factors (heritability). For several high heritability traits, it was possible to conduct genome-wide association studies (GWAS) and identify specific regions of the genome controlling between plant variation.

2 | MATERIALS AND METHOD

2.1 | Plant materials and image acquisition

Image data were acquired at the University of Nebraska Greenhouse Innovation Center's automated phenotyping facility (Ge et al., 2016) shown in Figure 1. Sorghum genotypes were taken from the Sorghum Association Population (Casa et al., 2008). The specific growth conditions for the sorghum plants grown and imaged in this study were previously described in the study by Miao, Pages, et al. (2020).

Each plant was photographed, and RGB images were collected from five side angles around a 360˚ with photos taken at 0˚, 72˚, 144˚, 216˚, and 288˚ plus one additional image from the top (see Figure 1d)). Each image had a resolution of 2,454 × 2,056 pixels. Plants were oriented so that the zero degree angle photo corresponded to the angle at which most leaves were perpendicular to a line between the camera and the primary stalk of the plant. The camera model is a Basler pia2400-17gc with a Pentax TV zoom lens c6z1218m3-5. Images used in this study were captured on April 11th, 2018, 47 days after planting. This dataset included 2,106 distinct images and was 17.5 GB in size.

Moreover, we also generated approximate 3D mesh models of corn plants using Plant Factory Exporter (v2016 R3, e-on Software) with the “maize/corn” module (Miao et al., 2019). The meshes were used for validation of the voxel carving algorithm by computing the precision and recall algorithm, as discussed in Section 3.

2.2 | Method

2.2.1 | 3D reconstruction overview

Our method works in four steps shown in Figure 2, where round boxes are processes and squared, boxes are data. The individual steps are: (a) data acquisition, (b) calibration, (c) segmentation, and (4) voxel carving. The input images are taken from the photographic chamber of the phenotyping facility, and the output of our algorithm is a 3D voxel representation of the plant. In this section, we provide an overview of our reconstruction pipeline explained in detail in the following sections.

Data acquisition

Is the process of taking images of a plant (see Section 2). The input is the physical plants, and the output is plant images. This process is fully automatic, each plant has an associated unique identifier, and the photos are stored on a local data server.

Calibration

To get an accurate reconstruction, we define a 3D coordinate system common to all images. A chosen reference point is the center of the pot that holds the plant. However, although the phenotyping facility attempts to align and center the pots on the rotational axis of the turntable when taking the photographs, in practice, they are usually off of the central axis by up to several centimeters. This causes the pot, and thus the plant, to be misaligned in the images. Since the alignment is one of the most critical condition for the voxel carving algorithm to work well, we calibrate the photos so that the pot is at the right location in every image. The calibration step results are a transformation matrix $\mathbf{T}$ that describes the translation that center the pot in each image.

Segmentation

In the next step, we separate the plant from the image background. Although the lighting conditions are controlled, the varying amount of plant mass and their geometry make the images difficult to segment using traditional methods, such as color segmentation or thresholding. Instead, we used a convolutional neural network that is invariant to changing light and works well with soft edges common in plants.

Voxel carving

Having the images segmented from the background and knowing the pose of the cameras, we reconstruct the plant using the voxel carving algorithm, which is a variant of the space carving algorithm of Kutulakos and Seitz (2000). The output is a 3D voxel grid representing the plant's photo hull, which is the maximal shape in which the actual plant lays. Although it is noticeably thicker than the plant, and does not include color information, it is still well suited for the heritability analysis (Section 3.4) and for other tasks such as counting the number of leaves and measuring their lengths.

Terminology

Let I denote the sequence of six input RGB images (see Figure 1d) $I = [I_0, I_{72}, I_{144}, I_{216}, I_{288}, I_{top}]$, where the first five images $[I_0, I_{72}, I_{144}, I_{216}, I_{288}]$ are taken from the sides with angles α in $A = [0˚, 72˚, 144˚, 216˚, 288˚]$ (see Figure 2a)), and I_{top} is the image taken from the top at an angle of 0˚. Moreover, $\mathbf{P} = [\mathbf{P}_0, \mathbf{P}_{72}, \mathbf{P}_{144}, \mathbf{P}_{216}, \mathbf{P}_{288}, \mathbf{P}_{top}]$ refer to the corresponding projection matrices of the camera while taking images I. To project a 3D point v on the camera with angle α,

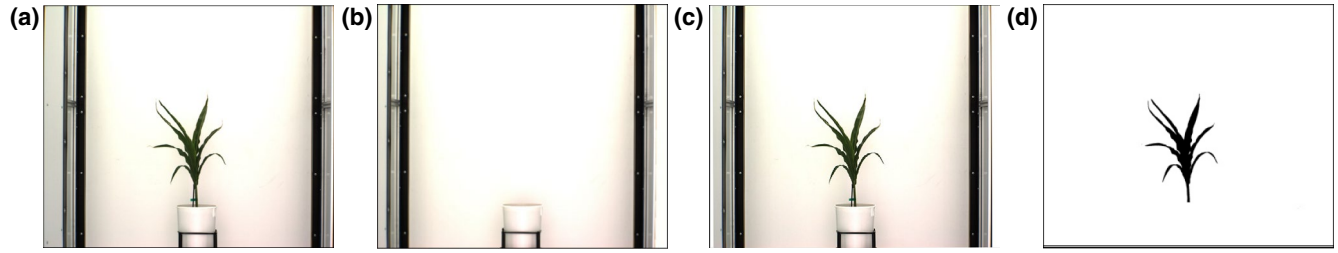

FIGURE 3 Image-based calibration and segmentation: (a) A raw image of a sorghum plant collected from the real imaging chamber I_0. (b) A centered reference image without a plant present created using the 3D virtual imaging chamber. (c) The output image which has been corrected and centered $\tilde{I}_0$. (d) The segmented mask generated from the image in panel C, with plant pixels indicated in black and not-plant pixels indicated in black. M_0

we multiply its coordinates by the projection matrix: $\tilde{v}_\alpha = \mathbf{P}_\alpha v$, where $\tilde{v}_\alpha$ is a 2D point in image I_α.

We denote the binary masks that result from the segmentation of I by $M = [M_0, M_{72}, M_{144}, M_{216}, M_{288}, M_{top}]$ (Figure 2f)). If the value of a pixel (k, l) in the mask $M_\alpha[k, l]$ is equal to one, the corresponding pixel in the image $\tilde{I}_\alpha[k, l]$ belongs to the object, while a value of zero indicates the projected position is in the background.

The final result of our algorithm is a binary voxel grid $\mathcal{V}$ that has a resolution of 512^3 and we refer to each voxel as $v_{i,j,k}$ where $0 \le i, j, k \le 512$ denote its discrete coordinates. A voxel $v_{i,j,k}$ having a value one identifies a 3D position belonging to the plant photo hull, zero value identifies an empty space.

2.2.2 | Image calibration

The phenotyping facility uses high precision robotic plant transportation conveyors shown in Figure 1a. Still, this precision is not sufficient to guarantee the plant to be perfectly centered in each image. Therefore, we use an image-based calibration method like in work from Roussel et al. (2016). The voxel carving algorithm is highly sensitive to the precise location, and thus we perform additional image transformations to make sure the plant is centered in every image I. The two primary sources of imprecision are (a) the pot is mounted on a turntable but is not perfectly centered on its axis of rotation, and (b) the camera's optical center is shifted from the axis of rotation of the turntable. This causes the plant to be off-axis when rotated.

To calibrate the images, we created a 3D virtual model of the phenotyping photographic chamber using a 3D modeling software. We then rendered a set of perfectly calibrated synthetic images of the virtual chamber without the plant, but with the pot at the exact center: one from the side and one from the top. These calibration images show how the pot should look in a theoretically perfect photographic chamber. We then used the images to find the translation that needs to be applied to the real images I to center the plant. We take a picture of the synthetic empty pot, and we found its location in the calibration image by overlaying it using transparency. Then, we look at the pot in each image I_α, and we use phase correlation first to shift the entire picture. Finally, we refine the transformation with template matching. This provides the transformation matrix $\mathbf{T}_\alpha$ that translates the image I_α so that the pot is centered. The corrected image set is denoted as $\tilde{I}$, and it is calculated by transforming the input image:

$$\tilde{I}_\alpha = \mathbf{T}_\alpha \times I_\alpha, \tag{1}$$

where $\times$ denotes the transformation of each pixel by the transformation matrix. In our dataset, we found that we had to shift images by up to 100 pixels horizontally to center the pot.

Our calibration only translates the images, and we did not need any other optical corrections. The sorghum plants are in the size range of meters, and the high-quality camera setup provides images with precision in the range of millimeters. The camera is located 5.5 m away from the plant, which lessens perspective distortion. Moreover, the camera has a high-quality optic, which limits other distortions, such as vignetting.

Figure 3a shows the input to the calibration step. The plant is not centered, and running voxel carving on this image would result in an imprecise reconstruction. Figure 3b shows the calibration image with an empty pot that is pixel-exact centered. Figure 3c shows the result of application of Equation (1) on I_0 where the input image is shifted so that the pot is at the same location as in the calibration image.

2.2.3 | Image segmentation

Although the imaging chamber is a controlled environment that removes a lot of variability from images, it cannot be directly inputted to the voxel carving algorithm because of the varying light intensity. While the chamber's lights are fixed and constant, the plant 3D structure and complex interreflections cause huge variability in the chamber's lighting. Therefore, it is beneficial to separate the plant in the image $\tilde{I}$ from the background.

Various techniques for image segmentation exists such as color thresholding (Lim & Lee, 1990) and color-based segmentation algorithms (Cheng, Jiang, Sun, & Wang, 2001; Haralick & Shapiro, 1985). However, as shown in Figure S5, they often struggle to segment the top view because they fail to differentiate between parts of the plant, which are in shadows cast by other parts of the plant and the

soil in the pot. Moreover, thresholding approaches tended to miss the leaf boundary in our hands, where the color gradient changes slowly.

After experimenting with an image processing software, we noticed that excellent results are generated manually by varying color mapping curves and applying different color-conversion filters. However, this work is tedious, cannot be automatized, and we could not find a fixed set of values that would work for all plants. Eventually, we used a convolutional neural network that was trained on human-segmented images. The machine learning approach takes into account a small neighborhood around each pixel for segmentation Adams et al. (2020). The training set contains 24 images, taken from four different plants with varying lighting conditions that were manually segmented. We used data augmentation to increase the variability of the dataset, particularly rotations of up to 90°, shift by up to 10% of the image size, zoom in by up to 10%, and horizontal and vertical flips. Having more extensive and diverse training data would doubtless aid in achieving better generalization. We found that this training dataset provided sufficient classification accuracy on the specific image data analyzed in this study for accurate downstream reconstruction.

The input to the neural network is a calibrated image $\tilde{I}_\alpha$, and the output is a binary image M_α of the segmented plant (Figure 3d). We use an architecture with four convolutional layers similar to the work of Donné et al. (2016). Our neural network (shown in Figure S1) does not include pooling layers, which means that our images are not downsampled during processing. When classifying a pixel, our network looks at a small neighborhood around it, applies non-linear operations, and outputs a probability of belonging to the background.

Similar to the work of Milletari, Navab, and Ahmadi (2016), we use the Dice loss for training because it is adapted to segmentation problems with imbalanced classes. The Dice loss is derived from the Dice similarity coefficient, commonly used for image segmentation validation. Having two shapes that are compared for similarity with the Dice coefficient, a value of 1 indicates perfect overlap, whereas a value of 0 indicates no similarity.

We split the dataset into 18 images for training and six images for validation. We trained for 2,000 epochs with the Adam optimization algorithm on a workstation equipped with a Xeon W-2145 (8 cores at 3.7GHz), 32 GB of RAM, and an Nvidia TITAN Xp with 12 GB of RAM. The training took about 3 hours, most of which spent on data augmentation.

2.2.4 | Voxel carving

The input to the voxel carving algorithm is the set of binary mask images M with a known camera projection matrix P. The output is a set of voxels that correspond to the plant. Our reconstruction algorithm is voxel-based volume carving that is a variant of the volume carving algorithm of Kutulakos and Seitz (2000). Our algorithm ignores color information in pixels (texture) and focuses only on geometry.

The plant is immersed into a uniform 3D volumetric grid denoted by $\mathcal{V}$ with a resolution 512^3. The grid has a physical size $1 \times 1 \times 1$ m^3 and at the resolution of 512^3 each voxel corresponds to a volume of about $2 \times 2 \times 2$ mm^3. As a reference, at the used distance from the camera, resolution, plant size and zoom level, each pixel in the side images represented an area of approximately 1.56 mm^2. This means that when projecting a voxel, it overlaps multiple pixels. Thus, we interpolate the pixel positions and take into account a small neighborhood around the projected voxel center when deciding whether to discard it. Note that the images capture an area slightly bigger than the voxel grid, so that its projection is strictly within all images from any point of view. Having a finer voxel grid would largely increase the processing time without bringing significantly more information.

A voxel center $v_{i,j,k} \in \mathcal{V}$ is projected on each mask image M_α using the projection matrix $\mathbf{P}_\alpha$. A voxel is set to one $v_{i,j,k} \leftarrow 1$ that is, it belongs to the plant, if the projected voxel hits pixel in each M_α that has also value set to one. If the voxel is projected onto at least one background pixel of the mask images (value zero), it is set to background that is, $v_{i,j,k} \leftarrow 0$:

$$v_{i,j,k} = \prod_{\alpha \in A} M_\alpha[\mathbf{P}_\alpha \times v_{i,j,k}]. \quad (2)$$

Figure 4 shows schematically this process for two masks rotated by 90^o and denoted by M_0 and M_{90}. The voxel $v_{i,j,k}$ is projected to M_0 using the projection matrix $\mathbf{P}_0$ and to M_{90} using $\mathbf{P}_{90}$. If the corresponding pixels in both masks are equal to one, the voxel $v_{i,j,k}$ is set to one.

The algorithm has a complexity of at least $\mathcal{O}(n^3)$ with n being the number of the voxels of the grid. However, this algorithm is also embarrassingly parallel, and it can process as many voxels in parallel as available processing units. Its output is a voxel grid $\mathcal{V}$ that associates to each voxel $v_{i,j,k}$ a binary value indicating whether it is a part of the

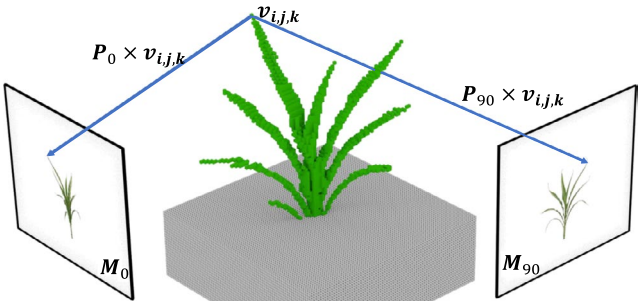

FIGURE 4 Illustration of the voxel carving algorithm employed in this study. The voxel carving algorithm projects each voxel $v_{i,j,k}$ to the corresponding masks (M_0 and M_{90} in this example) by multiplying its center by the calibrated projection matrix $\mathbf{P}$ indicated as rays. If the projected voxel's position in all masks corresponds to pixels with value one the voxel is also set as to one that is, as being a part of the plant

plant photo hull or not (the photo hull is the maximal shape that the plant occupies).

Voxel carving is highly sensitive to camera calibration. If one part of the plant is missing or distorted in only one image, it will not be included in the reconstructed volume (Equation 2). This causes problems with thin parts, such as the leaf tips or leaves projected from side.

To address this issue, we extended the voxel carving algorithm by looking at a small neighborhood in images instead of only a single pixel. When processing a voxel, we project a sphere of a fixed radius R on each input binary image M_α. If the projected sphere covers some pixels in the mask belonging to the plant, it is considered a match:

$$v_{i,j,k} \leftarrow 1 \quad \Leftrightarrow \quad \forall \alpha \in A, \exists V \in \mathcal{R}^3 \text{ s.t. } \|v_{i,j,k} - V\| \leq R \wedge M_\alpha[P_\alpha \times V] = 1 \tag{3}$$

This improves the recall of the reconstruction at the expense of precision as shown in the validation in Section 3.3. However, it is better to have higher recall values, because it corresponds to plant with continuous areas that are easier to post process. When using the sphere instead of a single value, the reconstructed plant is more faithful and includes less holes. The reconstructed leaves are thicker, but it better captures the overall plant shape. In other words, leaves are still at the correct location and their length and width are captured correctly as shown in Figure 5.

2.2.5 | Trait extraction and GWAS

The reconstructed voxels were used to quantitation of four traits for each sorghum line in SAP. The number of voxels in a reconstructed plant which approximates the plant volume is denoted by

$$vol(\mathcal{P}) = \sum_{v_{i,j,k} \neq 0} v_{i,j,k}. \tag{4}$$

where $vol(\mathcal{P})$ is the volume of a plant $\mathcal{P}$ represented by the total number of voxels. The bounding cylinder's volume was also calculated as an approximation of the space occupied by a plant. Each plant was encapsulated into its bounding cylinder, which approximates the space occupied by a plant. The bounding cylinder's volume is calculated in the cylindrical coordinate system, and it is tighter compared to the bounding box system. We constrain the axis of the bounding cylinder to be coincident with the z-axis.

Moreover, we also calculate the shadow area caused by light arriving from the top of the plant and it was denoted by $\text{shadow}(\mathcal{P})$ and calculated as

$$\text{shadow}(\mathcal{P}) = \sum_{v_{i,j,k} \neq 0} \mathbf{P}^{\perp}_{top} v_{i,j,k}, \tag{5}$$

where $\mathbf{P}^{\perp}_{\text{top}}$ is the top orthogonal projection matrix. If a projected voxel falls into shadow it is counted only once. The value of $\text{shadow}(\mathcal{P})$ is directly proportional to the cosine-corrected amount of intercepted light arriving from the top which approximates the amount of sun energy received by a plant (Benes, 1997; Soler, Sillion, Blaise, & Dereffye, 2003).

The ratio of the number of voxels to the shadow area indicates the light interception efficiency $\text{eff}(\mathcal{P})$ by a plant $\mathcal{P}$ and it is calculated as

$$\text{eff}(\mathcal{P}) = \frac{\text{shadow}(\mathcal{P})}{\text{vol}(\mathcal{P})}. \tag{6}$$

With a published set of 569,306 SNP markers for the same sorghum association population (Miao, Xu, et al., 2020) used in this study, the narrow-sense heritability was estimated for each above trait using GEMMA (Zhou & Stephens, 2012). GWAS analysis was also conducted for each trait to identify genes controlling the phenotypic variations within the population using the mixed linear model (MLM)-based GWAS algorithm implemented in GEMMA. The first

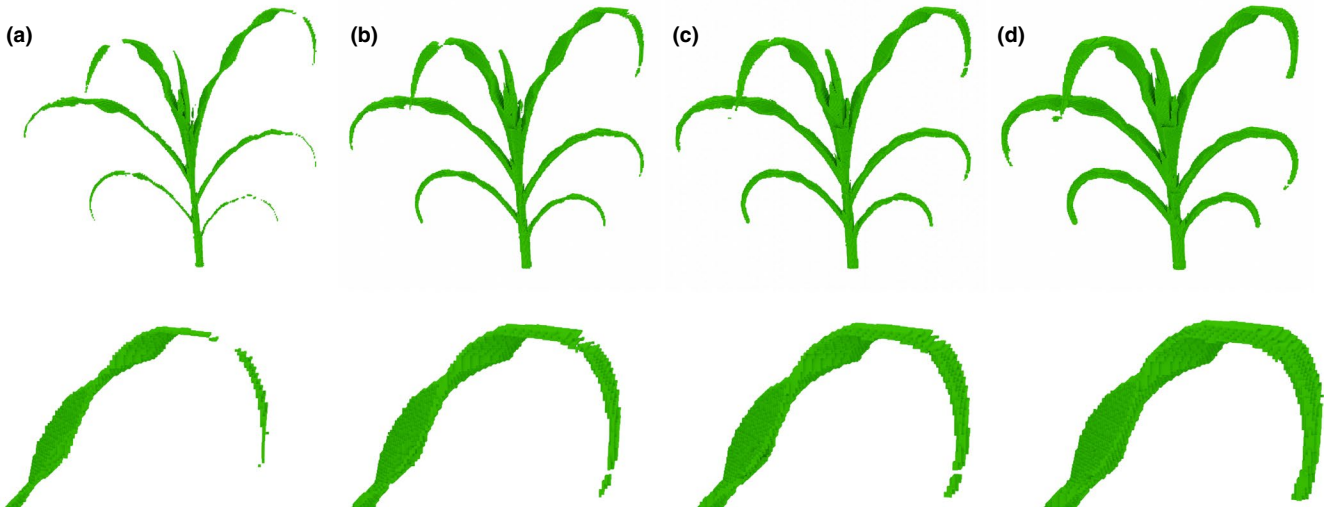

FIGURE 5 Effect of varying the size of the sensitivity area on the final 3D reconstruction of plant structures. Reconstructions of either an entire plant (top) or detailed view of the reconstruction of the upper right leaf from the same plant (bottom): when the sensitivity area is set to a sphere with a radius (a) one, (b) two, (c) three, and (d) four centimeters

three principal components calculated from Tassel (Bradbury et al., 2007) were fitted as fixed effects, and a kinship matrix calculated within GEMMA was fit as a random effect in the MLM. The number of independent SNPs estimated using the GEC/0.2 software package described by Li, Yeung, Cherny, and Sham (2012) were used to calculate the Bonferroni corrected *p*-value of .05 as the cutoff to determine statistically significant SNP-trait associations.

3 | RESULTS

3.1 | Segmentation validation

The average Dice similarity coefficient between the predicted validation set and the ground truth images was about 0.997 in our experiments. We visually inspected the masks and the segmentation in all photos. The result is correct except when the plant has a support that is sometimes classified as a background even if it is over the plant, which cuts the leaf in the reconstructed plant. In some top views, we also noticed that some dirt stains on the floor might be misclassified as belonging to the plant. However, having the wrong pixels only in one view is rejected in the voxel carving algorithm that uses the voting mechanism (Equation 3). An artifact that would be classified as a voxel would need to be in all views simultaneously that is virtually impossible. Although a deeper neural network could provide better results, we faced a memory limitation while training on our GPU. The images are large, which causes two problems: (a) the dataset is larger than the available RAM, and (b) we could not add many layers and filters. A possible avenue for future work is to training on the CPU with more RAM. Moreover, we could use the Tversky loss (Salehi, Erdogmus, & Gholipour, 2017), which is a generalization of the Dice loss, to favor recall over precision. When it comes to voxel carving, adding extra information in images is less problematic than removing essential information.

3.2 | 3D reconstruction validation

The accuracy of plant 3D reconstruction was assessed using synthetic data generated from procedural maize plants and an in silico reconstruction of the plant imaging chamber at the University of Nebraska-Lincoln using the dataset of Miao et al. (2019). We used a similar protocol as Roussel et al. (2016) and extended their work in the following ways: (a) we use a set of synthetic plants, which have a similar geometry to the plants we scan. (b) We compare different camera setups and show which give the best accuracy.

Ten triangle meshes of maize plants have been generated using Plant Factory Exporter (see e.g., in Figure S2). We visually inspected the 3D models to makes sure that they did not include self-intersections or other errors. The generated plants had 11 leaves on average. The minimum and maximum were 9 and 15, respectively. The ten meshes were voxelized, and the voxel grids were used as ground truth for 3D reconstruction.

We then simulated the data acquisition step by rendering six images per plant, five sides and one top view, by replicating the parameters employed for real-world data collection. These images were already calibrated and segmented since we tuned the simulation of the data acquisition to output directly the plant masks. Finally, we reconstructed the photo hull of the plants based on the six images and the six camera matrices using our voxel carving algorithm. Note that we used regular voxel carving, without our extension for poorly calibrated images. Because we generate perfect plant masks, we don't need to recover missing parts of the plant.

We compared the ground truth voxelized meshes with the ones reconstructed using our algorithm, and we use information retrieval evaluation measures: precision, recall, and F-measure. Precision is the fraction of reconstructed voxels that are part of the plant. Let's denote by *tp* (true positive) the number of detected voxels present in both grids—the ground truth and the reconstructed one. Let's denote by *fp* (false positive) the number of detected voxels present in the reconstructed grid but not in the ground truth, and let's denote *fn* (false negative) the number of detected voxels that are present in the ground truth, but were not reconstructed. We define precision *p* as

$$p = \frac{tp}{tp + fp}, \tag{7}$$

recall is denoted by *r*, and it is the fraction of voxels from the actual plant that were successfully reconstructed

$$r = \frac{tp}{tp + fn}, \tag{8}$$

and F-measure *Fm* combines precision and recall, and it is the harmonic mean of both measures

$$Fm = 2 \cdot \frac{p \cdot r}{p + r}. \tag{9}$$

We calculated *p*, *r*, and *Fm* for the 10 maize models, and we also calculated the average and standard deviation. This gives an upper bound on the accuracy we can get using voxel carving using our particular camera setup, and the results are in Table S3.

While precision $p \approx .5$, recall $r > .95$, which indicates that no essential parts of the plant are missing. As explained in Section 2.2.4, although the precision is not high, imprecision does not affect the overall shape of the plant and does not have a big impact on measurements such as plant height and leaf lengths. Figure 6 shows visual comparison of generated plants and their reconstruction. Furthermore, Figure 7a) shows a regression that predicts the true volume given the estimated volume for the 10 synthetic plants. The value of $R^2 = .84$ shows that counting the volume of voxels in the reconstructed plant is a valid estimation of the plant's true volume. The same approach applied to the estimation of the plant surface by counting the number of surface voxels is shown in Figure 7b), and it returned a value of $R^2 = .95$. Although we do not show the corresponding regressions in a figure, we

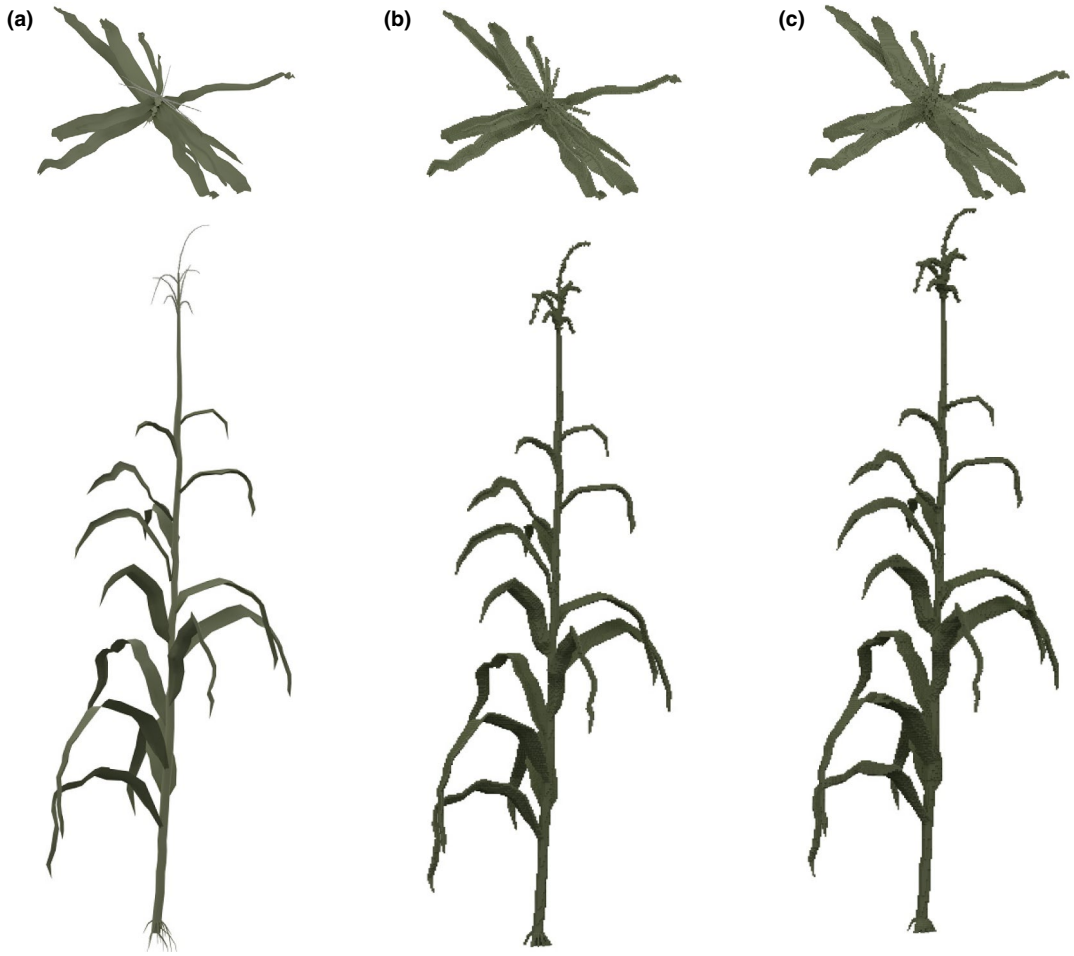

FIGURE 6 Estimating upper bound reconstruction accuracy using simulated data. Side and top views of (a) a procedurally generated 3D model of a maize plant, (b) a voxelized version of the original 3D model, and (c) a reconstructed 3D model generated using the voxel carving algorithm describes in this paper, and six simulated 2D images generated from the original 3D model (five side views and one top view)

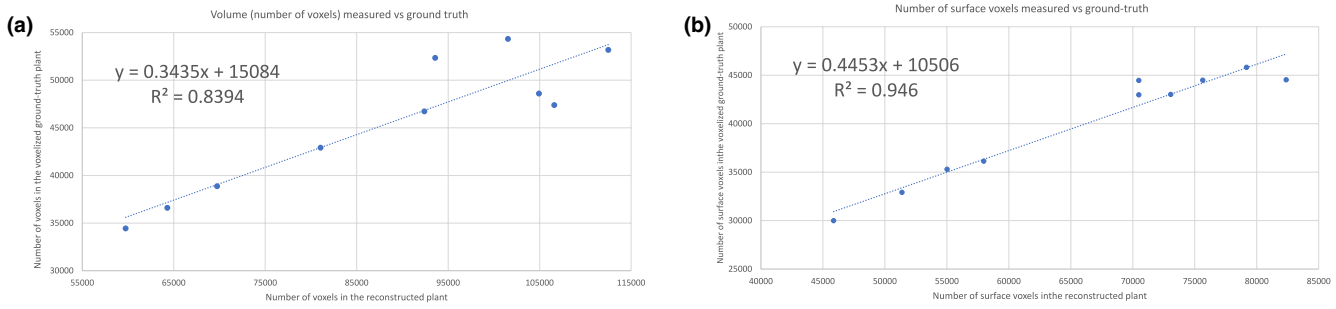

FIGURE 7 Correlation between ground truth and reconstructed plant properties. (a) The *y*-axis indicates the number of voxels present in a direct conversion of procedurally generated 3D models of maize plants into voxels. The *x*-axis indicates the number of voxels present in 3D models reconstructed using the approach to voxel carving used in this paper from 2D images generated from the same initial 3D models. A simple linear regression between the two values is shown, as is the Pearson coefficient of determination. (b) A similar analysis, instead of comparing the number of surface voxels in the direct voxel conversion of 3D models of corn plants (*y*-axis) and the number of surface voxels when instead reconstructing with voxel carving from 2D images

conducted the same analysis for other traits from Section 2.2.5: plant height, shadow area, the volume of the bounding cylinder and always found values of $R^2 \geq .99$. Section 3.5 discusses further details on how the camera setup affects the reconstruction accuracy.

3.3 | 3D reconstruction accuracy

We ran the reconstruction pipeline on hundreds of plants, and it would not be easy to verify each plant's accuracy manually. Inspired

by the work of Klodt and Cremers (2015), we computed the Dice coefficient denoted by D and used it to verify whether a plant was successfully reconstructed without visually inspecting it. We re-project all voxels from each plant to the six views, and we then compute D between the re-projected images and the plant masks obtained after the segmentation. The value of $D = 1$ indicates that everything in the plant masks has been reconstructed, and $D = 0$ indicates that the reconstruction failed (see Section 2.2.3 for details about the Dice coefficient). In our experiments, we found an average value of mean(D) = 0.884 with a standard deviation of stdev(D) = 0.011 for the 339 non-empty plants from our dataset.

To provide a visual insight into why some plants have a low Dice coefficient value, we color-coded the reprojections from each camera. We alpha-blended them with the corresponding RGB plant images. The results in Figure S3 show in green true positives, blue color encodes false positives, red value depicts false negative that is, a part of the plant that has not been reconstructed, and white color is a true negative.

We identified two factors that negatively affect the Dice coefficient. First, when taking pictures, leaf tips move as the plant rotates on the platform, and sometimes they are not entirely stable. Second, the segmentation step leaves small artifacts, especially on the top view, negatively impacting the Dice coefficient even if the reconstruction is accurate. Nevertheless, we can quickly identify a poorly reconstructed plant when its value is significantly off the distribution ($D < 0.8$ in our experiments).

We found that common errors (see e.g., in Figure S3) which caused the low value of Dice coefficient included (a) plants with incorrectly segmented dry leaves, (b) broken stems that make parts of the plant invisible from the top view or out of the voxel grid, (c) failed calibration due to leaves laying in the pot in one view, (d) small plants that cause noisy reconstruction, and (e) dead plants (empty pots).

3.4 | Variation in 3D structure among the sorghum association panel

We captured a dataset for 229 Sorghum lines in the sorghum association panel (SAP) at the early growth stage using a high throughput phenotyping facility (Ge et al., 2016; Miao et al., 2019) (Section 2.2). The dataset includes 1,374 (229 lines × 6 viewing angles) RGB images and an overall 12.0 GB. We ran the entire 3D construction pipeline on them.

The total surface area, which would be in shadow below each reconstructed sorghum plant, was calculated (Figure 8a). As expected, larger plants cast larger shadows, and this trait, which is referred to below as shadow area shadow($\mathcal{P}$) Equation (5), was significantly correlated with the total number of voxels in the plant vol($\mathcal{P}$)—a proxy for total plant volume (Figure 8b). The overall Pearson correlation coefficient of between plant voxel count and shadow area was 0.81 (p value = 2.8e-81). However, there was substantial variation in the ratio of shadow area to plant voxel count with the plants with the highest ratio casting approximately three times as much shadow per unit of volume as plants with the lowest ratios (Figure 8c). As shown in Figure 8d, large plants tend to be less light interception efficiency than the relatively smaller plants, but there is no clear linear relationship between these two features. For example, the sorghum line PI534105 casts a large shadow using a relatively small number of voxels. In contrast, PI533821 has a large number of voxels but casts a comparatively small shadow. After checking the 3D models of these two plants, we found that PI533821 has a more compact leaf architecture than PI534105 (see Figure 8b). The corresponding 3D files in OBJ format can be found in the Supplementary Materials.

Narrow-sense heritability was estimated for features extracted from 3D sorghum models. The estimated narrow-sense heritability for the number of voxels, bounding cylinder volume, shadow area, and the ratio (shadow area/number of voxels) representing light interception efficiency were 0.51, 0.49, 0.59, and 0.65 respectively. These results suggest that all the features above are sufficiently heritable for the mapping of individual loci controlling between genotype variation. GWAS analyses were conducted for each of the four phenotypes. Statistically significant trait associated SNPs were identified for both bounding cylinder volume representing space occupied by a plant and the ratio of shadow area to the number of voxels eff($\mathcal{P}$) representing light interception efficiency. Two significant signals were identified for bounding cylinder volume (Figure S4a,b). The significant peak on chromosome 7 likely corresponds to *dwarf3*,

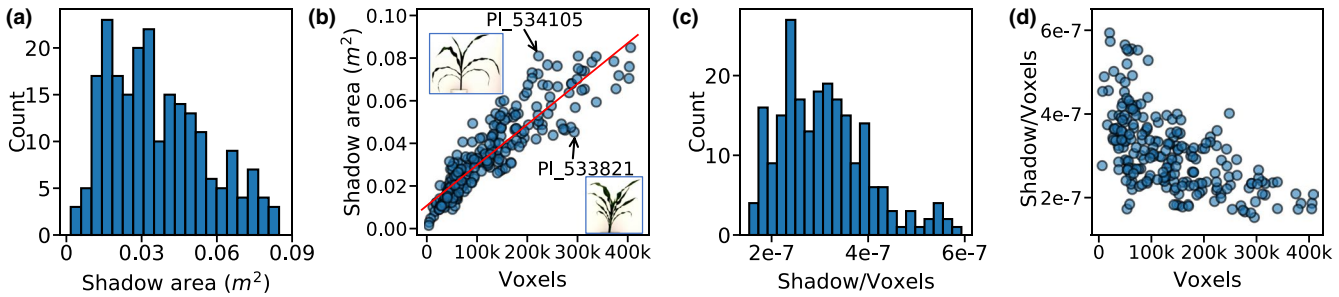

FIGURE 8 Distribution and relationship of radiation use efficiency related traits in the Sorghum Association Panel. (a) The distribution of shadow area shadow($\mathcal{P}$) Equation (5) across sorghum lines in the sorghum association panel (SAP) tested in this study. (b) Relationship between shadow area and the number of voxels. The best-fit linear regression line is indicated in red with the equation $y = 1.91e\text{-}7x + 0.01$. Two sorghum lines with particularly high or low ratios are indicated with arrows and their silhouettes are shown. (c) The distribution of ratio of shadow area to number of voxels across sorghum lines in the sorghum association panel (SAP) tested in this study. (d) Relationship between the shadow area/voxels ratio (e.g., eff($\mathcal{P}$) Equation (6) and the total number of voxels for a given plant

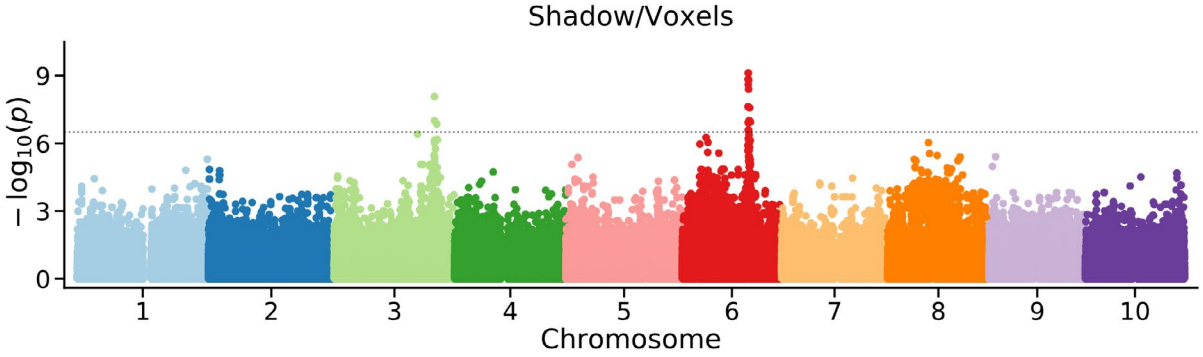

FIGURE 9 Genome-wide association analysis to identify genetic loci controlling variation of light interception efficiency in SAP. Manhattan plot summarizing the results of a genome-wide association study conducted using shadow area/voxels ratio representing light interception efficiency calculated from 3D reconstructions of sorghum lines from the SAP. Horizontal dashed line indicates a Bonferroni multiple testing-corrected threshold for statistical significance equivalent to p value = .05

a classical sorghum mutant that encodes an MDR transporter influencing cell elongation potentially via polar auxin transport (Multani et al., 2003). The second signal for bounding cylinder volume is a single SNP located within the gene Sobic.005G070200 on sorghum chromosome 5. Sobic.005G070200 encodes a wall-associated receptor kinase galacturonan-binding protein. Two well-supported and statistically significant clusters of trait associated SNPs were identified for light interception efficiency (shadow area/number of voxels) $\mathrm{eff}(\mathcal{P})$ (Figure 9). One of these peaks, on sorghum chromosome 6, corresponds to a second classical sorghum dwarf gene, *dwarf2* (Hilley et al., 2017). However, the second well-supported peak, on sorghum chromosome 3, is novel. The locations, p-values, and estimated effect sizes for the most significant SNP in each peak of each genome-wide association study presented here are provided in Supplementary Materials.

3.5 | Optimizing image acquisition for 3D reconstruction

Having a complete 3D model of the phenotyping chamber, we performed a set of virtual experiments that measure the reconstruction's behavior in varying conditions and result in several suggestions for image acquisition. We use the benchmark from Section 3.2 to evaluate different configurations of the camera setup.

We tested three parameters and checked how they affect the 3D reconstruction: (a) the number of captured images, (b) the presence of the top camera, and (c) the presence of an optional third camera looking at the plant from the middle of the angular distance between the top and the side cameras that is, at 45°. We call the configurations as *side* (side cameras), *top*, and *angle*. We compare the reconstruction accuracy when taking: 3, 5, 7, 9, 11, 13, and 15 *side* images, with or without the *angle* and *top* cameras. In total, we compared 28 different camera setups and results are shown in Figure 10 and the values in Table S1 in Supporting Material.

With the increasing number of captured images, the precision increases at the expense of recall independently of the presence of the *top* and the *side* camera. This behavior is expected: the number of detected voxels in voxel carving can only decrease when adding a new image. In other words, we reconstruct less volume of the plant, but what we get is more precise. The F-measure always increases as we add pictures. Taking more images improves the reconstruction accuracy. However, the marginal benefit—quantified by F-measure—for each additional side view image plateaus at relatively low values. Adding other cameras that observe the plant from different angles provides significant initial increases in precision and F-measure and reaches higher values before plateauing. Adding additional cameras to provide more viewing angles increases the cost of an imaging setup but not the time for data acquisition. In contrast, collecting more side views from the same camera does not increase the cost of construction, but significantly increases the time required per plant and decreases throughput. The system needs to wait for the plant to stabilize its motion after each rotation.

The *angle* and the *top* cameras always improve the reconstruction. Figure 10c) shows that with a fixed number of images adding a camera always yields to a better F-measure. An important observation is that the F-measure's effect is greater when adding the *angle* camera than the *top* camera. For the same investment (two cameras), we can get better data by setting up the second camera as *angle* instead of *top*, as is the common practice. Our phenotyping facility is a closed system that does not allow us to make modifications to the setup. However, Scharr et al. (2017) used a three-camera setup with a *angle* camera, and they report success in plant reconstruction.

The calibration step is essential for voxel carving, because the plant is not pixel-exact centered and the *side* camera always produced shifted images. We, thus, ran a study to estimate the loss in accuracy due to bad calibration. In particular, we run our reconstruction while simulating an imperfect calibration by translating simulated side images along the x-axis by 1 to 10 pixels. We report the F-measure of the reconstruction in Figure 10d) and in Table S2 in Supporting Material. A shift of 2 pixels caused a drop of 14% of the F-measure (from 0.6536 to 0.5638). Shifting the image by 10 pixels caused a loss of 91% (from 0.6536 to 0.0617).

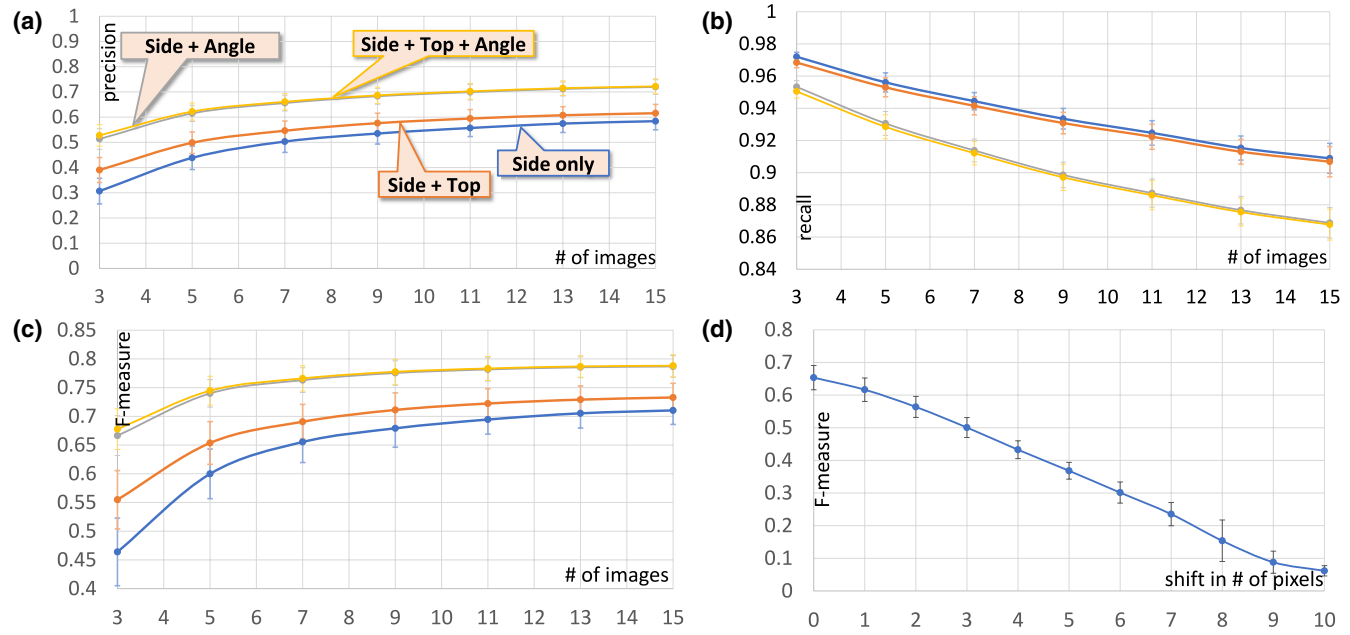

FIGURE 10 Impact of camera setup on reconstruction accuracy: (a) precision, (b) recall, and (c) F-measure with respect to the camera setup and the number of images taken around the plant. (d) F-measure with respect to the side image camera offset in pixels

4 | DISCUSSION

4.1 | Best practices when capturing images for future 3D reconstruction

Experiments in phenotyping facilities are expensive and time-consuming and generally cannot be easily repeated to change or improve image acquisition. In this section, based on our experience, we provide several guidelines to best use phenotyping facilities to maximize the amount of information captured and improve the potential for future reuse.

When scanning plants with long leaves, a good option is either the camera rotating or use multiple cameras as opposed to rotating the plant (Kumar, Connor, & Mikiavcic, 2014). We are not aware of a reconstruction algorithm that is robust enough to deal with leaves that move between pictures. It is better to put the camera farther from and change the focal length so that it occupies most of the image, rather than placing it close. For a plant occupying the same area in an image, a remote camera with a long focal length flattens perspective distortion, reduces blur, and maximizes the per-pixel precision of images relative to a close camera with a short focal length. In Section 3.5, we simulated and evaluated a range of imaging setups and camera position options. Adding more images improves the outcome, but only up to a certain point due to a plateau effect. An alternate way to enhance the reconstruction accuracy is to add more camera angles. When only one camera is available, we recommend taking side views with it. If a second camera is available, our virtual experiments show that a high-angle shot provides more information than a camera looking directly down on the plant.

Colors in RGB images have considerable importance in allowing or hindering the segmentation of plant pixels from background pixels. A white background makes it easy to separate a green plant. When using supports for the plants, such as pots or a duct tape, it is best to avoid green or dark colors because they are much more likely to be misclassified than bright pixels such as blue, red, or yellow. The main reason is that when using a segmentation method based on colors, it is hard to distinguish green and dark colors from the plant, as shaded parts of the plant can be closer to black than green.

When using voxel carving with only a few images—a common scenario in most phenotyping facilities—the voxel reconstruction will often contain artifacts that look like leaves. Many of these can be quickly discarded by selecting the major connected component of voxels, as many artifacts will not be connected to the real plant. However, we still observed some artifacts attached to plants. These prove more challenging to remove, and future work is needed to address them. In the short term, we can only urge researchers to collect data using automated phenotyping facilities to capture as many views from as many distinct viewing angles as the logistics of their experimental design allows.

4.2 | GWAS analysis on traits from reconstructed 3D plants

Here we identified genetic loci in sorghum controlling variation in traits: (a) the ratio of total plant volume to the size of the shadow cast by a plant, a trait we have referred to as light interception efficiency, and (b) volume of the bounding cylinder required to contain a plant. Both of these traits would be challenging to qualify either

from a single 2D image or through conventional manual measurements in the field or greenhouse. For each trait, two significant loci were identified. In each case, one of these loci corresponded to a gene with a known role in controlling variation of other plant traits, and the other had, to our knowledge, not previously been identified in GWAS for other sorghum traits.

In particular, light interception efficiency may be a useful test case of the value of assessing the properties of different genotypes of the same species from 3D reconstructions. It is possible to quantitate the total leaf area of individual plants through destructive sampling. However, in the absence of information on the positioning of these leaves in 3D space, two plants with the same total leaf area and height may intercept very different quantities of overall light if in one plant lower leaves are in the direct shadow of upper leaves and the second. In this study, thanks to the constructed 3D plant, we can estimate the light interception efficiency defined as the normalized plant shadow area by the total number of plant voxels $\mathrm{eff}(\mathcal{P})$ Equation (6). The estimated narrow-sense heritability of 0.65 indicates the light interception efficiency was controlled mainly by the genetic factor, and the GWAS results showed two significant peaks. The first peak is at chromosome 3, which had not been reported in any studies related to canopy architecture or leaf morphology. The second considerable peak is close to one of the classical dwarf genes dw2. Previous research showed that dw2 not only controls plant height but also has a pleiotropic effect on leaf area (Graham & Lessman, 1966; Pereira & Lee, 1995).

However, it is essential not to confuse the property of light interception efficiency with radiation use efficiency. Radiation use efficiency is commonly defined as the quantity of dry matter produced per unit of radiant energy captured. This depends not only on total light energy hitting the surface of the leaf, which can be estimated from our 3D models plus weather data, date, and latitude, but also on the efficiency of photosynthesis, carbon uptake and transpiration, etc. However, changes in canopy architecture can indeed result in changes in radiation use efficiency. For example, changes in leaf angle that distribute high-intensity light over a larger area of photosynthetic tissue, increasing the proportion of light which can be put to productive use somewhat being lost when the photosynthetic apparatus is overwhelmed. Changes in light distribution in the canopy increase the balance of photosynthesis happening deep in the canopy where water loss due to transpiration is reduced can also increase radiation use efficiency.

The GWAS for bounding cylinder volume identified two significant peaks located in chromosomes 5 and 7, respectively (Figure S4). The first peak only contains one SNP within the gene body of Sobic.005G070200, a sorghum gene encoding a cell wall-associated receptor kinase. The second peak is a cluster containing seven continuous SNPs close to one of the cloned dwarf genes—dw3. Dw3 influences sorghum internode length and the leaf angle—two essential traits in determining the plant occupied space. Truong et al. had shown dw3 as a major leaf angle QTL with the recessive allele decreases leaf inclination angle up to 34° (Truong, McCormick, Rooney, & Mullet, 2015). At least three large-effect genes for plant height are known to segregate in the population. However, previous reports have indicated that detecting these genes at intermediate stages of development can be challenging (Miao, Xu, et al., 2020). While it was also possible to quickly estimate plant height from the 3D reconstructed plants, and, as described above, two of these genes were identified in GWAS for 3D traits, no significant associations were identified. This suggests that the pleiotropic effects of the sorghum dwarfing genes may be more natural to detect across development than the direct impact of these genes on plant height.

4.3 | Limitations

A limitation inherent to voxel carving is that it cannot reconstruct cavities or voids that are not observable from the outside. Here we employ the algorithm to reconstruct individual sorghum plants without tillers (secondary stalks growing from the same base). In this context, the algorithm works acceptably well. We anticipate that it would translate well to plants with similar architectures such as a maize, foxtail millet, and pearl millet. However, for plants with highly branched architectures where leaves obscure their interior structure from all angles (e.g., soybeans or tomatoes), voxel carving is unlikely to provide satisfactory results.

Voxel carving also works best for extremely rigid plants or plants photographed from multiple angles simultaneously rather than sequentially. Here plants were photographed using a turntable and a single side camera. As a result of the rotation and air movements, leaves are unlikely to be in the same positions between photos, necessitating the neighborhood approach employed in this study. The severity of this problem will be the least for rigid plants such as woody deciduous perennials when they don't have leaves, and the greatest for highly flexible plants such as vines. The best way to mitigate this issue is to use a network of fixed cameras and no turntable. If cameras are triggered at the same time, the plants can't move between shots. While this leads to a more difficult calibration process, a practical solution has been proposed in (Tabb, Medeiros, Feldmann, & Santos, 2019). Our current solution results in over-voxelization. While Section 3.2 demonstrates that over-voxelization does not present a series of issues for the traits described in Section 2.2.5, it is anticipated that our present method would not work well for estimating some other traits such as leaf thickness.

Finally, we computed light interception on plants in isolation and from only a single angle, rather than taking into account the movement of the sun through the sky. Thus, we did not consider the potential effects of neighboring plants on each other in the field and how different canopy architectures may serve to distribute light in different ways through the canopy as the sun moves through the sky. This initial study employed greenhouse-grown plants that likely exhibit different architectures from the same genotypes grown at agriculturally relevant planting densities under field conditions. Future studies could use the same algorithms defined here to quantitate the 3D architectures of field-grown plants and use these models to model entire plant canopies, providing

estimates of the quantity of light intercepted, the amount of biomass volume required to intercept that light, and how light was distributed throughout different parts of the canopy under field conditions.

ACKNOWLEDGMENTS

We thank Melba Crawford and Behrokh Nazeri for discussion and initial help with the data retrieval. Valérian Méline for discussion about plant biology and limits of current phenotyping and Lydia Lindner for helping successfully training the segmentation neural network. This project was completed utilizing the Holland Computing Center of the University of Nebraska, which receives support from the Nebraska Research Initiative.

CONFLICT OF INTEREST

The authors are not aware of any conflict of interest arising from drafting this manuscript.

DATA AVAILABILITY STATEMENT

The authors commit to depositing both the raw images analyzed as part of this study, voxel-based 3D reconstructions, and trait values for 3D phenotypes quantified for each plant into DataDryad once the manuscript is accepted for publication. Genetic marker data employed in this study was previously published in Miao, Pages, et al. (2020) and has been deposited in FigShare DOI: 10.6084/m9.figshare.11462469.v5. The code is available at https://github.com/cropsinsilico/SorghumVoxelCarving.

ORCID

Mathieu Gaillard https://orcid.org/0000-0003-0840-5654
Chenyong Miao https://orcid.org/0000-0002-0904-3707
James C. Schnable https://orcid.org/0000-0001-6739-5527
Bedrich Benes https://orcid.org/0000-0002-5293-2112

REFERENCES

AdamsJ., Qiu Y., Xu Y., Schnable J. C. (2020). Plant segmentation by supervised machine learning methods. *The Plant Phenome Journal*, *3*, e20001. http://doi.org/10.1002/ppj2.20001.

Alexandratos, N., & Bruinsma, J. (2012). World agriculture towards 2030/2050: The 2012 revision. The 2012 Revision Ch. 4 (ESA/12-03, FAO, 2012). ESA Working Papers 12-03. https://doi.org/10.22004/ag.econ.288998

Benes, B. (1997) Visual Simulation of plant development with respect to influence of light. In D. Thalmann & M. de Panne (Eds.), *Computer animation and simulation'97* (Springer Computer Science, pp. 125–136). Vienna, Austria: Springer-Verlag.

Benes, B., Guan, K., Lang, M., Long, S. P., Lynch, J., Marshall-Colon, A., Peng, B., Schnable, J. C., Sweetlove, L., & Turk, M. (2020). Multiscale computational models can guide experimentation and targeted measurements for crop improvement. *The Plant Journal*, *103*, (1), 21–31. http://dx.doi.org/10.1111/tpj.14722.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). Tassel: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*, 2633–2635.

Cabrera-Bosquet, L., Fournier, C., Brichet, N., Welcker, C., Suard, B., & Tardieu, F. (2016). High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform. *New Phytologist*, *212*, 269–281.

Casa, A. M., Pressoir, G., Brown, P. J., Mitchell, S. E., Rooney, W. L., Tuinstra, M. R., … Kresovich, S. (2008). Community resources and strategies for association mapping in sorghum. *Crop Science*, *48*, 30–40.

Cheng, H.-D., Jiang, X. H., Sun, Y., & Wang, J. (2001). Color image segmentation: Advances and prospects. *Pattern Recognition*, *34*, 2259–2281.

Das ChoudhuryS., Samal A., Awada T. (2019). Leveraging image analysis for high-throughput plant phenotyping. *Frontiers in Plant Science*, *10*. http://doi.org/10.3389/fpls.2019.00508

Donné, S., Luong, H., Goossens, B., Dhondt, S., Wuyts, N., Inzé, D., … Philips, W. (2016). Machine learning for maize plant segmentation. In *Belgian-Dutch Conference on Machine Learning (BENELEARN)*, Kortrijk, Belgium.

Du, S., Lindenbergh, R., Ledoux, H., Stoter, J., & Nan, L. (2019). Adtree: Accurate, detailed, and automatic modelling of laserscanned trees. *Remote Sensing*, *11*, 2074.

Duvick, D. N. (2005). The contribution of breeding to yield advances in maize (zea mays l). *Advances in Agronomy*, *86*, 83–145.

Fahlgren, N., Feldman, M., Gehan, M. A., Wilson, M. S., Shyu, C., Bryant, D. W., … Ficor, T., (2015). A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in setaria. *Molecular Plant*, *8*, 1520–1535.

Gaillard, M., Miao, C., Schnable, J., & Benes, B. (2020). Sorghum segmentation by skeleton extraction. In *ECCV 2020 Workshops* (to appear).

Ge, Y., Bai, G., Stoerger, V., & Schnable, J. C. (2016). Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Computers and Electronics in Agriculture*, *127*, 625–632.

Gehan, M. A., Fahlgren, N., Abbasi, A., Berry, J. C., Callen, S. T., Chavez, L., … Hoyer, J. S., (2017). Plantcv v2: Image analysis software for high-throughput plant phenotyping. *PeerJ*, *5*, e4088.

Gibbs, J. A., Pound, M., French, A. P., Wells, D. M., Murchie, E., & Pridmore, T. (2017). Approaches to three-dimensional reconstruction of plant shoot topology and geometry. *Functional Plant Biology*, *44*, 62–75.

Golbach, F., Kootstra, G., Damjanovic, S., Otten, G., & Zedde, R. (2016). Validation of plant part measurements using a 3d reconstruction method suitable for high-throughput seedling phenotyping. *Machine Vision and Applications*, *27*, 663–680.

Graham, D., & Lessman, K. (1966). Effect of height on yield and yield components of two isogenic lines of sorghum vulgare pers. 1. *Crop Science*, *6*, 372–374.

Grassini, P., Eskridge, K. M., & Cassman, K. G. (2013). Distinguishing between yield advances and yield plateaus in historical crop production trends. *Nature communications*, *4*, 1–11.

Hammer, G. L., Dong, Z., McLean, G., Doherty, A., Messina, C., Schussler, J., … Cooper, M. (2009). Can changes in canopy and/or root system architecture explain historical maize yield trends in the us corn belt? *Crop Science*, *49*, 299–312.

Haralick, R. M., & Shapiro, L. G. (1985). Image segmentation techniques. *Computer Vision, Graphics, and Image Process*, *29*, 100–132.

Hilley, J. L., Weers, B. D., Truong, S. K., McCormick, R. F., Mattison, A. J., McKinley, B. A., … Mullet, J. E. (2017). Sorghum dw2 encodes a protein kinase regulator of stem internode length. *Scientific Reports*, *7*, 4616.

Junker, A., Muraya, M. M., Weigelt-Fischer, K., Arana-Ceballos, F., Klukas, C., Melchinger, A. E., … Altmann, T. (2015). Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. *Frontiers in Plant Science*, *5*, 770.

Klodt, M., & Cremers, D. (2015). High-resolution plant shape measurements from multi-view stereo reconstruction. In *Computer Vision—ECCV 2014Workshops* (pp. 174–184). Cham, Switzerland: Springer International Publishing.

Koenderink, N., Wigham, M., Golbach, F., Otten, G., Gerlich, R., & van de Zedde, H. (2009). Marvin: High speed 3D imaging for seedling classification. In *Precision agriculture'09* (pp. 279–286). Wageningen, The Netherlands.

Kumar, P., Connor, J., & Mikiavcic, S. (2014). High-throughput 3D reconstruction of plant shoots for phenotyping. In *2014 13th International Conference on Control Automation Robotics Vision (ICARCV)*, Singapore (pp. 211–216).

Kutulakos, K. N., & Seitz, S. M. (2000). A theory of shape by space carving. *International Journal of Computer Vision*, *38*, 199–218.

Li, B., Heng, L., Koser, K., & Pollefeys, M. (2013). A multiple-camera system calibration toolbox using a feature descriptorbased calibration pattern. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Tokyo, Japan. (pp. 1301–1307).

Li, M.-X., Yeung, J. M., Cherny, S. S., & Sham, P. C. (2012). Evaluating the effective numbers of independent tests and significant *p*-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics*, *131*, 747–756.

Lim, Y. W., & Lee, S. U. (1990). On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques. *Pattern Recognition*, *23*, 935–952.

Lobet, G. (2017). Image analysis in plant sciences: publish then perish. *Trends in Plant Science*, *22*, 559–566.

Lou, L., Liu, Y., Sheng, M., Han, J., & Doonan, J. H. (2014). A cost-effective automatic 3d reconstruction pipeline for plants using multi-view images. In M. Mistry, A. Leonardis, M. Witkowski, & C. Melhuish (Eds.), *Advances in autonomous robotics systems* (pp. 221–230). Cham, Switzerland: Springer International Publishing.

Maddonni, G., Chelle, M., Drouet, J.-L., & Andrieu, B. (2001). Light interception of contrasting azimuth canopies under square and rectangular plant spatial distributions: Simulations and crop measurements. *Field Crops Research*, *70*, 1–13.

Marshall-Colon, A., Long, S. P., Allen, D. K., Allen, G., Beard, D. A., Benes, B., … Hirst, P. M., (2017). Crops in silico: Generating virtual crops using an integrative and multi-scale modeling platform. *Frontiers in Plant Science*, *8*, 786.

McCormick, R. F., Truong, S. K., & Mullet, J. E. (2016). 3D sorghum reconstructions from depth images identify QTL regulating shoot architecture. *Plant Physiology*, *172*, 823–834.

MiaoC., Hoban, T. P., Pages A., Xu Z., Rodene E., Ubbens J., … Schnable J. C. (2020). Semantic Segmentation of Sorghum Using Hyperspectral Data Identifies Genetic Associations. *bioRxiv*. http://dx.doi.org/10.1101/706994

MiaoC., Pages A., Xu Z., Rodene E., Yang J., & Schnable J. C. (2020). Semantic Segmentation of Sorghum Using Hyperspectral Data Identifies Genetic Associations. *Plant Phenomics*, *2020*, 1–11. http://dx.doi.org/10.34133/2020/4216373

MiaoC., Xu Y., Liu S., Schnable P. S., & Schnable J. (2020). Increased power and accuracy of causal locus identification in time-series genome-wide association in sorghum. *Plant Physiology*, pp.00277.2020http://dx.doi.org/10.1104/pp.20.00277

Milletari, F., Navab, N. and Ahmadi, S.-A. (2016) V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)* (pp. 565–571). Stanford, California: IEEE.

Multani, D. S., Briggs, S. P., Chamberlin, M. A., Blakeslee, J. J., Murphy, A. S., & Johal, G. S. (2003). Loss of an MDR transporter in compact stalks of maize br2 and sorghum dw3 mutants. *Science*, *302*, 81–84.

Paulus, S., Schumann, H., Kuhlmann, H., & Léon, J. (2014). High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants. *Biosystems Engineering*, *121*, 1–11.

Pendleton, J., Smith, G., Winter, S., & Johnston, T. (1968). Field investigations of the relationships of leaf angle in corn (zea mays l.) to grain yield and apparent photosynthesis 1. *Agronomy Journal*, *60*, 422–424.

Pepper, G., Pearce, R., & Mock, J. (1977). Leaf orientation and yield of maize 1. *Crop Science*, *17*, 883–886.

Pereira, M., & Lee, M. (1995). Identification of genomic regions affecting plant height in sorghum and maize. *Theoretical and Applied Genetics*, *90*, 380–388.

Quan, L., Tan, P., Zeng, G., Yuan, L., Wang, J., & Kang, S. B. (2006). Image-based plant modeling. *ACM Transactions on Graphics*, *25*, 599–604. http://doi.acm.org/10.1145/1141911/1141929.

QuinbyJ. R., Karper R. E. (1954). Inheritance of Height in Sorghum1. *Agronomy Journal*, *46*, (5), 211.http://doi.org/10.2134/agronj1954.00021962004600050007x.

Roussel, J., Geiger, F., Fischbach, A., Jahnke, S., & Scharr, H. (2016). 3D surface reconstruction of plant seeds by volume carving: Performance and accuracies. *Frontiers in Plant Science*, *7*, 745.

Salehi, S. S. M., Erdogmus, D., & Gholipour, A. (2017) Tversky loss function for image segmentation using 3D fully convolutional deep networks. In *International Workshop on Machine Learning in Medical Imaging* (pp. 379–387). Quebec City, Canada: Springer.

Scharr, H., Briese, C., Embgenbroich, P., Fischbach, A., Fiorani, F., & Müller-Linow, M. (2017). Fast high resolution volume carving for 3D plant shoot reconstruction. *Frontiers in Plant Science*, *8*, 1680. https://www.frontiersin.org/article/10.33.89/fpls.2017.01680.

Soler, C., Sillion, F. X., Blaise, F., & Dereffye, P. (2003). An efficient instantiation algorithm for simulating radiant energy transfer in plant models. *ACM Transactions on Graphics*, *22*, 204–233. https://doi.org/10.1145/636886.636890.

St'ava, O., Pirk, S., Kratt, J., Chen, B., Měch, R., Deussen, O., & Benes, B. (2014). Inverse procedural modelling of trees. *Computer Graphics Forum*, *33*, 118–131. https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.12282.

Tabb, A., Medeiros, H., Feldmann, M. J., & Santos, T. T. (2019). Calibration of asynchronous camera networks: Calico. arXiv preprint arXiv:1903.06811.

Tan, P., Fang, T., Xiao, J., Zhao, P., & Quan, L. (2008). Single image tree modeling. *ACM Transactions on Graphics*, *27*, 108:1–108:7. http://doi.acm.org/10.1145/1409060/1409061

Tan, P., Zeng, G., Wang, J., Kang, S. B., & Quan, L. (2007). Image-based tree modeling. *ACM Transactions on Graphics*, *26*, 87–es. http://doi.acm.org/10.1145/1276377.1276486

Thapa, S., Zhu, F., Walia, H., Yu, H., & Ge, Y. (2018). A novel lidar-based instrument for high-throughput, 3D measurement of morphological traits in maize and sorghum. *Sensors*, *18*, 1187.

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, *108*, 20260–20264.

Tomasi, C., & Kanade, T. (1992). Shape and motion from image streams under orthography: A factorization method. *International Journal of Computer Vision*, *9*, 137–154.

Truong, S. K., McCormick, R. F., Rooney, W. L., & Mullet, J. E. (2015). Harnessing genetic variation in leaf angle to increase productivity of sorghum bicolor. *Genetics*, *201*, 1229–1238.

Ward, B., Bastian, J., van den Hengel, A., Pooley, D., Bari, R., Berger, B., & Tester, M. (2015). A model-based approach to recovering the structure of a plant from images. In L. Agapito, M. M. Bronstein, & C. Rother (Eds.), *Computer Vision—ECCV 2014 Workshops* (pp. 215–230). Cham, Switzerland: Springer International Publishing.

Westgate, M., Forcella, F., Reicosky, D., & Somsen, J. (1997). Rapid canopy closure for maize production in the northern us corn belt: Radiation-use efficiency and grain yield. *Field Crops Research*, *49*, 249–258.

YangWanneng, Guo Zilong, Huang Chenglong, Duan Lingfeng, Chen Guoxing, Jiang Ni, Fang Wei, Feng Hui, Xie Weibo, Lian Xingming, Wang Gongwei, Luo Qingming, Zhang Qifa, Liu Qian, Xiong Lizhong (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature Communications*, *5*. http://doi.org/10.1038/ncomms6087.

Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, *44*, 821.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Gaillard M, Miao C, Schnable JC, Benes B. Voxel carving-based 3D reconstruction of sorghum identifies genetic determinants of light interception efficiency. *Plant Direct*. 2020;00:1–16. https://doi.org/10.1002/pld3.255